Coarse-Grained Modeling of Genetic Circuits as a Function of the Inherent Time Scales

Darka Labavić[1], Hannes Nagel[2], Wolfhard Janke[2] and Hildegard Meyer-Ortmanns[1]

[1]School of Engineering and Science, Jacobs University Bremen,
P.O.Box 750561, 28725 Bremen, Germany
[2]Institut für Theoretische Physik, Universität Leipzig,
Postfach 100 920, 04009 Leipzig, Germany

E-mail: `d.labavic@jacobs-university.de`, `hannes.nagel@itp.uni-leipzig.de`,
`wolfhard.janke@itp.uni-leipzig.de` `h.ortmanns@jacobs-university.de`

Abstract. From a coarse-grained perspective the motif of a self-activating species, activating a second species which acts as its own repressor, is widely found in biological systems, in particular in genetic systems with inherent oscillatory behavior. Here we consider a specific realization of this motif as a genetic circuit, in which genes are described as directly producing proteins, leaving out the intermediate step of mRNA production. We focus on the effect that inherent time scales on the underlying fine-grained scale can have on the bifurcation patterns on a coarser scale in time. Time scales are set by the binding and unbinding rates of the transcription factors to the promoter regions of the genes. Depending on the ratio of these rates to the decay times of the proteins, the appropriate averaging procedure for obtaining a coarse-grained description changes and leads to sets of deterministic equations, which differ in their bifurcation structure. In particular the desired intermediate range of regular limit cycles fades away when the binding rates of genes are of the same order or less than the decay time of at least one of the proteins. Our analysis illustrates that the common topology of the widely found motif alone does not necessarily imply universal features in the dynamics.

1. Introduction

A frequently found motif in biological networks, in particular in genetic networks, is the combination of a positive feedback loop in which one species (A) activates itself, and a negative feedback loop, in which the first species activates its own repressor, the second species (B). In connection with genetic systems such motifs are realized in the cAMP signalling system in the slime mold Dictyosthelium Discoideum [1], in the embryonic division control system [2, 3, 4], the MAPK-cascade [5], or in the circadian clock [6, 7]. From the viewpoint of theoretical physics one is interested in dynamical features that are in common to the many different dynamical realizations of this motif. One common feature is certainly the occurrence of regular oscillations and the possibility of excitable behavior for an appropriate choice of parameters, and these features are captured already by a deterministic description in form of the bistable frustrated unit, considered in [8] and references therein. It should be emphasized, however, that the different realizations do not only differ by the biological systems in which they are realized, but also by the degree in which this representation as two coupled loops as shown in Fig. 1 amounts
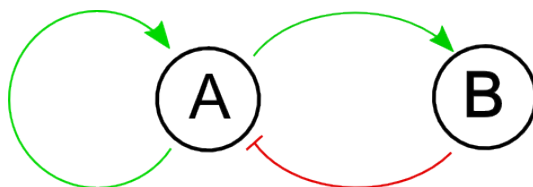

Figure 1. Basic motif of a self-activating species A, activating also its own repressor B. Pointed arrows denote activation, blunt arrow denotes repression.

only to an effective description. The effective description should be seen in contrast to a one-to-one mapping of the participating ingredients. In principle a number of intermediate steps may be included in these loops, and these intermediate steps need not be on the same level in case of a hierarchical organization. They could amount to reactions between genes leading to production rates on the level of proteins, or on the same level, just to reactions between proteins, and it may make a big difference if protein A is repressed via the binding of a transcription factor of type B to gene b (design I), or via a direct repression of protein A via B (design II), as it was emphasized in [9]. There it was shown that the distinct set of equations, corresponding to the two designs, lead to different oscillatory features and different behavior with respect to noise and external periodic signals, so that the motif in the first design acts as an "integrator" of external stimuli, in the second design as a "resonator".

For a coarse-grained description that should reflect the essential dynamical implications of this motif, the question therefore arises when it is actually justified to subsume the intermediate steps to effectively effective single steps and to represent the whole system by a set of two ordinary differential equations for the concentrations of A and B. In particular such a description implies some kind of large-volume limit, in which the stochastic fluctuations of demographic origin, that is in the number of individuals, are

ignored. In [8] we have analyzed the effect of demographic fluctuations and of fluctuations in the reaction times in simple realizations of this motif, that is, without any intermediate steps. The stochastic fluctuations in the number of reactants and in the reaction times led to the occurrence of so-called quasi-cycles in parameter regimes for which the system would be deeply in the fixed-point regimes in the deterministic limit. On the other hand, the three regimes of the deterministic limit could be clearly recognized: as function of one bifurcation parameter α, the maximal rate of production of A for full activation and no repression, the dynamics of the coupled species A and B converges to a fixed point and shows excitable behavior in a first phase. In a second phase, regular limit-cycle behavior is observed, and in a third phase, again a fixed point with excitable behavior is seen. The quasi-cycles which are only due to the finite system size, are more regular in the transition regions than far off in the fixed point phases, they can be distinguished from regular oscillations by the decay of their autocorrelation functions.

When this motif is supposed to describe a genetic circuit, a possible zoom into the dynamical details would amount to a description in terms of proteins A and B, their associated mRNA at an intermediate level and two types of genes a and b, respectively. Protein production of type A would result from transcription factors of type A, activating gene a which is transcribed to the corresponding mRNA that leads to the translation to proteins A. At the same time, protein A leads to an activation of gene b via binding of the transcription factor A to gene b, which is then transcribed to mRNA of type b, and when translated to proteins leads to a repression of the production of A.

In general, such intermediate steps as the production of mRNA and the binding and unbinding of transcription factors to the promoter region of genes induce additional time scales: delay of the protein production, and the switching rate between gene states. Here we distinguish the following gene states: states with bound activating transcription factor (the gene being in the "on-state"), states with bound repressing transcription factor (the gene being in the "off-state"), or no transcription factor bound (which we call the gene being in the "bare-state"). In principle we could consider further situations such as the simultaneous binding of an activator and a repressor to the promoter region of the gene a, but we leave out this option for simplicity. Whether the delay time and the switching rate can be ignored compared to the time scales defined by the protein decay rates of A and B is a matter of relative size. In addition, the excitable behavior with large excursions in phase space of the original motif is related to the pronounced difference in the decay times of the proteins: A is considered as the fast variable and B as the slow variable, similarly to the distinction of fast and slow variables in a FitzHugh-Nagumo element [10, 11]. (Due to this inherent difference, the oscillations become more spiky, the excursions in phase space from a fixed point more pronounced, and the description more realistic if we think of neurons realizing these circuits.) Therefore, when comparing the switching rate of gene states, we should specify with respect to which decay rate (that of A or of B) it is fast or slow. In the following we call the switching rates "fast" if they are fast in comparison to both decay rates, of A and of B, "slow" if the switching rate is of the order of the fast decay rate of A, and "ultra-slow" if it is of the order of

the slow decay rate of B.

In former related studies of this genetic circuit [9] the switching rates of genes were assumed to be so high that their effect could be assumed to average out in the sense that mRNA and proteins see only average values of gene activation or repression. In this paper we want to analyze the effect that the inherent time scale of binding and unbinding rates of transcription factors has on the coarse-grained modeling. In order to project on this effect, we neglect the intermediate step of mRNA production, but vary the binding/unbinding rates to values which are no longer high, but of the same order as the decay times of either the "fast" protein A or of the "slow" protein B.

From results for genetic switches, in particular for the toggle switch [12], a simpler system than ours with only two mutually repressing genes, it is known that additional fixed points may show up for (what we call) slow genes. More precisely, the joint probability distribution of the number of proteins in the cell and the DNA-binding site state (being on or off) have additional peaks, corresponding to additional fixed points in the deterministic limit, if the switches are studied in the nonadiabatic limit. If the switching rate between the genes is slow enough, the remainder of the system can follow the corresponding gene state of being either on or off. The dynamics of our system is more versatile. In addition to the fixed-point regime in the deterministic limit, we have the regime of regular limit-cycles, and the decay time of proteins differs by a factor of the order of hundred. So independently of the possibility to realize our system in synthetic genetic circuits or to find it in natural systems, we want to focus on the impact of "slow" and "ultra-slow" components on the system's dynamics, in particular on the phase structure.

We start from a fully stochastic description in terms of master equations describing reactions directly between genes and proteins, skipping the intermediate mRNA-production steps. We simulate these equations via the Gillespie algorithm [13]. In view of a coarse-grained description in the form of deterministic equations we take appropriate averages of the master equations which are adapted to the various limiting cases. These averaging procedures correspond to taking the zeroth order of a van Kampen expansion [14] of the master equations, where the expansion parameter is the inverse system size. The phase structure then differs for the resulting sets of deterministic equations, it depends on the competing time scales of genes and proteins. For fast genes we recover the previously observed structure of three regimes, two fixed-point and one limit-cycle regime. For this case we provide a detailed bifurcation analysis. For slow and ultra-slow genes, the regime of regular limit cycles is absent. Instead, the fast proteins A see (certain combinations of) distinct states of gene expression separately rather than their averages: in the deterministic limit they therefore approach the fixed points corresponding to these states and stay there, unless the gene states change; in the stochastic realization they switch between these states, so that the probability density to find N_A proteins of type A and N_B proteins of type B shows as many peaks as there are gene states distinguished. In contrast, the inherent dynamics of the proteins B is still so slow that these proteins cannot reach the vicinity of these fixed points in the stochastic

realization which they would approach in the deterministic description, leading to broad peaks or almost uniform probability densities.

The subsequent sections of the paper are organized as follows. In sect. 2 we present the model of interacting genes and proteins for the motif of a positive self-activating feedback loop coupled to a negative feedback loop with a second species repressing the first one. The model is described in terms of chemical reactions and master equations. In sect. 3 we derive deterministic models along with a stability analysis and a comparison with Gillespie simulations for three limiting cases: fast genes (sect. 3.1), slow genes (sect. 3.2) and ultra-slow genes (sect. 3.3). The summary and conclusions are given in sect. 4, followed by an Appendix in which we analyze in detail the two transition regions between fixed-point behavior and limit cycles.

2. The model

In terms of biochemical reactions we consider the following realization of the motif of Fig. 1 displayed in Fig. 2.

Proteins A and B are produced under different conditions on the expression level of genes, but we assume in all cases that the production is proportional to the system size. The system size is parameterized by a factor N_0. For protein A we distinguish three
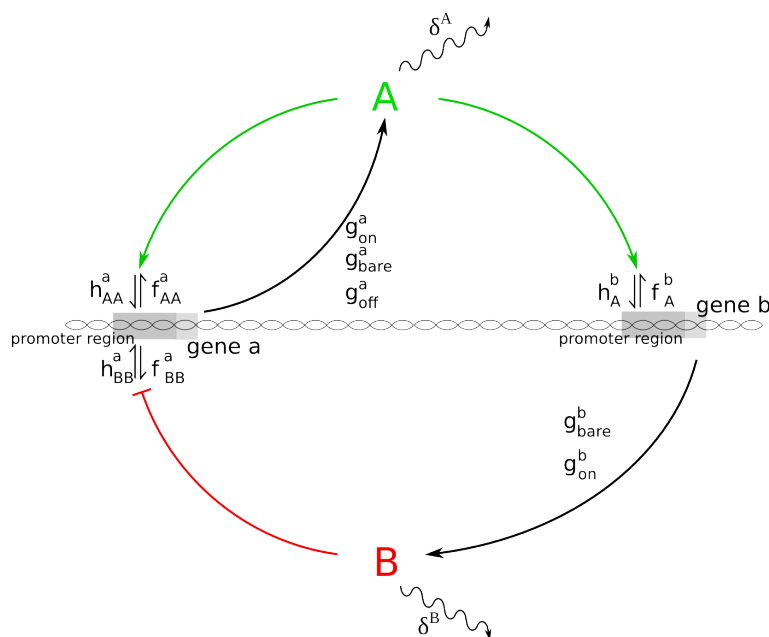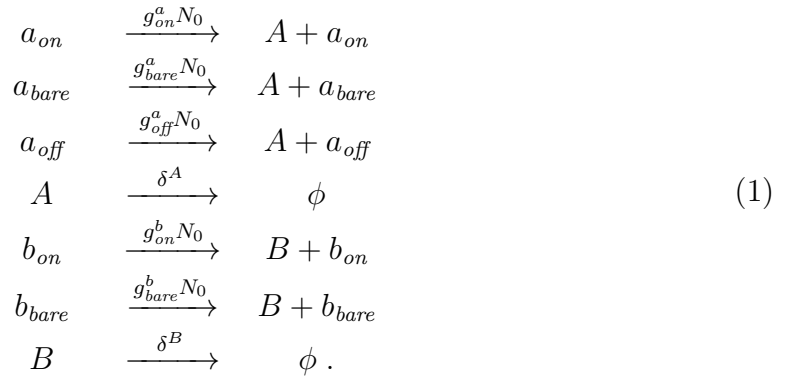

Figure 2. Zoom into the motif of Fig. 1 with a realization via genes a and b leading to the production of proteins A and B with rates $g^a_{on,bare,off}$ and $g^b_{bare,on}$, respectively, depending on the bound transcription factors to the promoter region of a and b. Transcription factors $A(B)$ bind with rate $h^a_{AA}(h^a_{BB})$ to the promoter region of a and unbind with rates $f^a_{AA}(f^a_{BB})$, respectively. Transcription factor A also binds to the promoter region of b with rate h^a_A and unbinds with rate f^b_A. Proteins A and B decay with rates δ^A and δ^B, respectively. For further explanations see the main text.
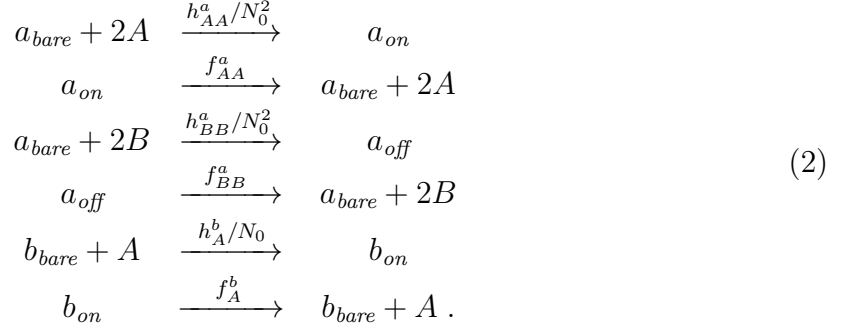
situations: (i) An activating transcription factor A is bound to the promoter region of gene a, so that gene a is in the on-state and produces proteins A accordingly with rate $g_{on}^a \cdot N_0$; here the superscript a and subscript on indicate that gene a is responsible for the production of prbbotein A and is itself in the on-state due to the binding of the (self-)activating transcription factor A. (ii) No transcription factor is bound to the promoter region of gene a, leading to a production of A in the so-called bare state of the gene with rate $g_{bare}^a \cdot N_0$. (iii) A repressing transcription factor B is bound to the promoter region of gene a and turns the gene to the lower expression level, so that we term the gene state to be "off" and the production rate of protein A proceeds with rate $g_{off}^a \cdot N_0$, accordingly. For simplicity we leave out a further possible situation that an activating and a repressing transcription factor A and B, respectively, are simultaneously bound to the respective promoter regions of gene a, leading to a conflicting input of activation and repression.

Since protein B is only activated, but not repressed via A, we distinguish here only two states: (i) An activating transcription factor A is bound to the promoter region of gene b, leading to a production of B with rate $g_{on}^b \cdot N_0$, or, (ii) no transcription factor is bound, leading to a production rate of $g_{bare}^b \cdot N_0$ of protein B. Moreover, protein A decays with rate δ^A and protein B with rate δ^B. Choosing δ^B much smaller than δ^A and $g_{on,bare}^b$ much smaller than $g_{on,bare}^a$ is a way to implement the different inherent time scales in the protein dynamics: that of A is much faster than that of B, the reason for why we call A the fast variable and B the slow variable also in this realization of the motif. The decay rate of the fast protein A sets our time scale, $\delta^A = 1$. Throughout this paper we choose $\delta^B = 0.01$, so that the slow protein B lives by a factor of 100 longer than A. The reactions referring to production and decay of proteins are summarized in the following equations:

$$
\begin{aligned}
a_{on} &\xrightarrow{\;g_{on}^a N_0\;} & A + a_{on} \\
a_{bare} &\xrightarrow{\;g_{bare}^a N_0\;} & A + a_{bare} \\
a_{off} &\xrightarrow{\;g_{off}^a N_0\;} & A + a_{off} \\
A &\xrightarrow{\;\delta^A\;} & \phi \\
b_{on} &\xrightarrow{\;g_{on}^b N_0\;} & B + b_{on} \\
b_{bare} &\xrightarrow{\;g_{bare}^b N_0\;} & B + b_{bare} \\
B &\xrightarrow{\;\delta^B\;} & \phi \, .
\end{aligned}
\tag{1}
$$

Next we come to the binding/unbinding reactions of transcription factors to the promoter region of genes. We distinguish between dimer binding of A and B to gene a (corresponding to a Hill coefficient of 2) and monomer binding of A to the promoter region of gene b (corresponding to a Hill coefficient of 1). (The Hill coefficient provides a quantitative measure for the binding cooperativity.) The corresponding binding reactions

along with the unbinding ones are listed in Eq. (2).

$$
\begin{aligned}
a_{bare} + 2A &\xrightarrow{h^a_{AA}/N_0^2} & a_{on} \\
a_{on} &\xrightarrow{f^a_{AA}} & a_{bare} + 2A \\
a_{bare} + 2B &\xrightarrow{h^a_{BB}/N_0^2} & a_{off} \\
a_{off} &\xrightarrow{f^a_{BB}} & a_{bare} + 2B \\
b_{bare} + A &\xrightarrow{h^b_A/N_0} & b_{on} \\
b_{on} &\xrightarrow{f^b_A} & b_{bare} + A\,.
\end{aligned}
\tag{2}
$$

The notation is chosen as follows: The superscript in the binding/unbinding coefficients h^i, f^j indicate the gene whose promoter region is affected in the binding/unbinding event, the subscripts refer to the monomer (one index) or dimer (two indices) binding/unbinding of the transcription factors. In our Gillespie simulations we choose the effective binding rates and the corresponding unbinding rates to be of the same order such that

$$
\frac{h^a_{AA}N_A^2}{N_0^2} = \frac{h^a_{BB}N_B^2}{N_0^2} = \frac{h^b_A N_A}{N_0} \sim f^a_{AA} = f^a_{BB} = f^b_A\,.
\tag{3}
$$

We normalize the binding rates which are proportional to the number of proteins N_A with the appropriate power in N_0 to make them approximately independent of the system size, assuming that $N_0 \sim N_A$. Our simulations have shown that in case of a ratio of binding and unbinding rates different from one, it is the smaller value of the binding and unbinding rates that determines the dynamics in the following sense: The implied changes affect only the location of the fixed points and the limit-cycle regime, but lead to no qualitative change.

Given now the common values for the binding and unbinding rates of transcription factors, we call the switching of gene states, induced by the binding and unbinding events, fast if these rates are much higher than the decay rate of the fast protein (set to one), slow if it is of the order of the fast protein A, and ultra-slow if it is of the order of the slow protein B.

As we shall argue in sect. 3.1, the A-protein production rate in the on-state, g^a_{on}, is chosen as the bifurcation parameter, while the production rates in the other states of gene a are kept fixed, such that g^a_{bare} is by an order of magnitude smaller than the usual values of g^a_{on}, and g^a_{off} is set to zero. The production rates of the protein B are also kept fixed and chosen by two orders of magnitude smaller than the corresponding production rates of protein A to compensate for the hundred times longer lifetime of protein B in the competing gain and loss terms in Eq. (25), see Sect. 3.1. Our choice of parameters is summarized in Tables 1 and 2.

The reactions described by Eqs.(1) and (2) correspond to a set of master equations which tell us the change in time of the probability to find at time t N_A proteins of type A and N_B proteins of type B, given that gene a is in either of the three states $i = on, bare, off$ and at the same time gene b is either in the on-state $j = on$, or in the bare-state

Table 1. Fixed parameters.

g^a_{bare}	g^a_{off}	g^b_{on}	g^b_{bare}	δ^A	δ^B
25	0	2.5	0.025	1	0.01

Table 2. Binding and unbinding parameters.

genes	N_0	N_A	N_B	$f^a_{AA} = f^a_{BB} = f^b_A$	$\frac{h^a_{AA}}{N_0^2}N_A^2 = \frac{h^a_{BB}}{N_0^2}N_B^2$	$\frac{h^b_A}{N_0}N_A$	$\delta^A \gg \delta^B$
fast	1	100	100	100	100	100	$\gg \delta^A \gg \delta^B$
slow	1	100	100	1	1	1	$\sim \delta^A \gg \delta^B$
ultra-slow	1	100	100	0.01	0.01	0.01	$\sim \delta^B \ll \delta^A$

$j = bare$. This probability is denoted as $P_{ij}(N_A, N_B, t)$. The six master equations for $P_{ij}(N_A, N_B, t)$, resulting from the six combinations of indices, can be summarized in matrix notation according to:

$$\frac{dP_{i,j}(N_A, N_B; t)}{dt} = -\left(g^a_i N_0 + \delta^A N_A + g^b_j N_0 + \delta^B N_B\right) P_{i,j}(N_A, N_B, t)$$

$$+ g^a_i N_0 P_{i,j}(N_A - 1, N_B, t) + \delta^A(N_A + 1) P_{i,j}(N_A + 1, N_B, t)$$

$$+ g^b_j N_0 P_{i,j}(N_A, N_B - 1, t) + \delta^B(N_B + 1) P_{i,j}(N_A, N_B + 1, t)$$

$$+ h^a_{AA}(i)\frac{N_A^2}{N_0^2} P_{bare,j}(N_A, N_B, t) - f^a_{AA}(i) P_{on,j}(N_A, N_B, t)$$

$$+ h^a_{BB}(i)\frac{N_B^2}{N_0^2} P_{bare,j}(N_A, N_B, t) - f^a_{BB}(i) P_{off,j}(N_A, N_B, t)$$

$$+ h^b_A(j)\frac{N_A}{N_0} P_{i,bare}(N_A, N_B, t) - f^b_A(j) P_{i,on}(N_A, N_B, t)$$

$$i = on,\ bare,\ off,\ j = on,\ bare, \tag{4}$$

if we introduce the following definitions:

$$h^a_{AA}(on) = -h^a_{AA}(bare) \equiv h^a_{AA},$$
$$h^a_{BB}(off) = -h^a_{BB}(bare) \equiv h^a_{BB},$$
$$h^a_{AA}(off) = h^a_{BB}(on) = 0 \tag{5}$$

for the dimer binding factors, and

$$f^a_{AA}(on) = -f^a_{AA}(bare) \equiv f^a_{AA},$$
$$f^a_{BB}(off) = -f^a_{BB}(bare) \equiv f^a_{BB},$$
$$f^a_{AA}(off) = f^a_{BB}(on) = 0 \tag{6}$$

for the dimer unbinding, where the argument in the bracket refers to the state of either gene a or gene b, referred to via i or j in the master equations, respectively. For monomer binding/unbinding factors we define

$$h^b_A(on) = -h^b_A(bare) \equiv h^b_A,$$
$$f^b_A(on) = -f^b_A(bare) \equiv f^b_A. \tag{7}$$

The first eight terms of the master equation for $P_{ij}(N_A, N_B, t)$ for all values of i, j result from the production or decay of proteins, leading to gain or loss terms as follows: Loss terms contributing to the change of $P_{ij}(N_A, N_B, t)$ are due to the production of protein A with constant rate $g_i^a \cdot N_0$, or the decay of A proportional to N_A, and the production of protein B with rate $g_j^b \cdot N_0$, or the decay of B proportional to N_B. Gain terms, on the other hand, are due to the production of A from a state with $N_A - 1$ proteins A, and the production of B from a state with $N_B - 1$ proteins B, or the decay of one protein A from a state with $N_A + 1$ proteins A, or the decay of one protein B from a state with $N_B + 1$ proteins B.

The last six terms in Eq. (4), of which between four and six are different from zero, describe changes in the probability due to the binding and unbinding of proteins to the promoter regions of gene a and b. For example, a positive contribution to the probability $P_{on,j}(N_A, N_B, t)$ results from a binding of a dimer of proteins A with rate $h_{AA}^a(on)\frac{N_A^2}{N_0^2}$, given that the system before the binding event has N_A, N_B proteins with gene a being in the bare state; a negative contribution results from an unbinding of a dimer of A-proteins with rate $f_{AA}^a(on) = f_{AA}^a$, given that the system contains N_A, N_B proteins while gene a being in the on-state before the unbinding event. The other terms are derived similarly.

3. Coarse-grained description of the genetic circuit

The coarse-graining that we intend to achieve refers primarily to a coarse-graining in time, averaging over events on a time scale much shorter than the scale on which the effective description should hold. (Usually, along with the coarse-graining in time goes a coarse-graining in space, as it is familiar from other realms of physics, but this is not in our focus here, as we do not arrange the genetic circuit and its constituents in any spatial ordering. Here the coarse-graining in time leads in general to a reduction of variables that are needed to describe the system's dynamics, but the amount of reduction depends on the limits which are taken, as we shall see below.) In order to derive coarse-grained descriptions in the form of deterministic equations for the protein concentrations $\Phi_A := N_A/N_0$ and $\Phi_B := N_B/N_0$, we apply appropriate averaging procedures to the master equation (4). We assume that the states of the genes are independent of each other and of the number of proteins N_A and N_B. Therefore we factorize the probability $P_{ij}(N_A, N_B, t)$ according to

$$P_{ij}(N_A, N_B, t) \ = \ a_i b_j \, P(N_A, N_B, t), \tag{8}$$

with a_i the probability of finding gene a in the i-state and b_j the probability of finding gene b in the j-state, while $P(N_A, N_B, t)$ is the probability of finding the respective

protein numbers whatever states the genes are in. We then consider expectation values

$$
\begin{aligned}
\langle N_S \rangle_{ij} &\equiv \sum_{N_A, N_B} P_{ij}(N_A, N_B, t) \, N_S \\
&= \sum_{N_A, N_B} a_i b_j P(N_A, N_B, t) \, N_S \\
&= \langle N_S \rangle a_i b_j \,,
\end{aligned} \tag{9}
$$

where S stands for the species A or B and the average represents a summation over all N_A, N_B values. We have to postulate

$$
\sum_i a_i = 1 = \sum_j b_j = \sum_{N_A, N_B} P(N_A, N_B, t), \tag{10}
$$

so that

$$
a_i = \sum_{N_A, N_B, j} P_{ij}(N_A, N_B, t), \quad b_j = \sum_{N_A, N_B, i} P_{ij}(N_A, N_B, t) \tag{11}
$$

and

$$
\sum_{N_A N_B} \frac{dP_{ij}(N_A, N_B)}{dt} = \frac{d(a_i b_j)}{dt}, \quad \sum_j \frac{d(a_i b_j)}{dt} = \frac{da_i}{dt}, \quad \sum_i \frac{d(a_i b_j)}{dt} = \frac{db_j}{dt}. \tag{12}
$$

Moreover, using (12) and (9), we have

$$
\sum_{N_A N_B} \frac{dP_{ij}(N_A, N_B, t)}{dt} N_S = \frac{d(\langle N_S \rangle a_i b_j)}{dt} \,; \qquad \sum_{i,j} \langle N_S \rangle a_i b_j = \langle N_S \rangle \tag{13}
$$

and

$$
\sum_{i,j} \frac{d(\langle N_S \rangle a_i b_j)}{dt} = \frac{d\langle N_S \rangle}{dt}. \tag{14}
$$

Now we multiply the master equation (4) with N_A and N_B, respectively, and sum over all N_A, N_B; next we sum the average value of N_A, $\langle N_A \rangle$ only over all states of gene b, since the change in N_A is assumed to be independent on the states of gene b, and in analogy $\langle N_B \rangle$ only over all states of gene a, respectively. We then obtain

$$
\frac{d}{dt} \left(\sum_{N_A, N_B, j} N_A P_{ij}(N_A, N_B, t) \right) = \frac{d}{dt} (\langle N_A \rangle a_i), \tag{15}
$$

where

$$
\frac{d\left(\langle N_A \rangle a_{on}\right)}{dt} = g_{on}^a N_0 a_{on} + h_{AA}^a \frac{\langle N_A \rangle^3}{N_0^2} a_{bare} - f_{AA}^a \langle N_A \rangle a_{on} - \delta^A \langle N_A \rangle a_{on} \tag{16}
$$

$$
\begin{aligned}
\frac{d(\langle N_A \rangle a_{off})}{dt} = {}& g_{off}^a N_0 a_{off} + h_{BB}^a \langle N_A \rangle \frac{\langle N_B \rangle^2}{N_0^2} a_{bare} - f_{BB}^a \langle N_A \rangle a_{off} \\
& - \delta^A \langle N_A \rangle a_{off}
\end{aligned} \tag{17}
$$

$$
\begin{aligned}
\frac{d\left(\langle N_A \rangle a_{bare}\right)}{dt} = {}& g_{bare}^a N_0 a_{bare} - \delta^A \langle N_A \rangle a_{bare} - h_{AA}^a \frac{\langle N_A \rangle^3}{N_0^2} a_{bare} + f_{AA}^a \langle N_A \rangle a_{on} \\
& - h_{BB}^a \langle N_A \rangle \frac{\langle N_B \rangle^2}{N_0^2} a_{bare} + f_{BB}^a \langle N_A \rangle a_{off} \,.
\end{aligned} \tag{18}
$$

Similarly,

$$\frac{d}{dt}\left(\sum_{N_A,N_B,i} N_B P_{ij}(N_A, N_B, t)\right) = \frac{d}{dt}(\langle N_B\rangle b_j), \tag{19}$$

where

$$\frac{d(\langle N_B\rangle b_{on})}{dt} = g_{on}^b N_0 b_{on} - \delta^B \langle N_B\rangle b_{on} + h_A^b \frac{\langle N_A\rangle}{N_0}\langle N_B\rangle b_{bare} - f_A^b \langle N_B\rangle b_{on} \tag{20}$$

$$\frac{d(\langle N_B\rangle b_{bare})}{dt} = g_{bare}^b N_0 b_{bare} - \delta^B \langle N_B\rangle b_{bare} - h_A^b \frac{\langle N_A\rangle}{N_0}\langle N_B\rangle b_{bare} + f_A^b \langle N_B\rangle b_{on}. \tag{21}$$

From Eq. (11) we obtain

$$\frac{da_{on}}{dt} = h_{AA}^a \frac{\langle N_A\rangle^2}{N_0}a_{bare} - f_{AA}^a a_{on}$$

$$\frac{da_{bare}}{dt} = -h_{AA}^a \frac{\langle N_A\rangle^2}{N_0}a_{bare} + f_{AA}^a a_{on} - h_{BB}^a \frac{\langle N_B^2\rangle}{N_0^2}a_{bare} + f_{BB}^a a_{off}$$

$$\frac{da_{off}}{dt} = h_{BB}^a \frac{\langle N_B\rangle^2}{N_0^2}a_{bare} - f_{BB}^a a_{off}$$

$$\frac{db_{on}}{dt} = h_A^b \frac{\langle N_A\rangle}{N_0}b_{bare} - f_A^b b_{on}$$

$$\frac{db_{bare}}{dt} = -h_A^b \frac{\langle N_A\rangle}{N_0}b_{bare} + f_A^b b_{on}. \tag{22}$$

These equations determine the time dependence of the probability to find the system in any of the five different gene states. In a stationary state of the genes, the left-hand side of Eqs. (22) vanishes. This leads to

$$a_{on} = \frac{\frac{h_{AA}^a}{f_{AA}^a}\frac{\langle N_A\rangle^2}{N_0^2}}{1 + \frac{h_{AA}^a}{f_{AA}^a}\frac{\langle N_A\rangle^2}{N_0^2} + \frac{h_{BB}^a}{f_{BB}^a}\frac{\langle N_B\rangle^2}{N_0^2}}$$

$$a_{bare} = \frac{1}{1 + \frac{h_{AA}^a}{f_{AA}^a}\frac{\langle N_A\rangle^2}{N_0^2} + \frac{h_{BB}^a}{f_{BB}^a}\frac{\langle N_B\rangle^2}{N_0^2}}$$

$$a_{off} = \frac{\frac{h_{BB}^a}{f_{BB}^a}\frac{\langle N_B\rangle^2}{N_0^2}}{1 + \frac{h_{AA}^a}{f_{AA}^a}\frac{\langle N_A\rangle^2}{N_0^2} + \frac{h_{BB}^a}{f_{BB}^a}\frac{\langle N_B\rangle^2}{N_0^2}}$$

$$b_{on} = \frac{\frac{h_A^b}{f_A^b}\frac{\langle N_A\rangle}{N_0}}{1 + \frac{h_A^b}{f_A^b}\frac{\langle N_A\rangle}{N_0}}$$

$$b_{bare} = \frac{1}{1 + \frac{h_A^b}{f_A^b}\frac{\langle N_A\rangle}{N_0}}, \tag{23}$$

using $\sum_i a_i = 1 = \sum_j b_j$. If the genes change their state fast enough as compared to other time scales in the system, they will reach the stationary values of Eq. (23) before the other processes are completed; therefore they may be used in equations like (16–21).

Now we are prepared to discuss the different limiting cases of fast, slow and ultra-slow genes.

3.1. Fast genes

In the limit of fast genes we choose all binding and unbinding rates a hundred times larger than the decay rate of the fast protein A, that is δ^A. In this limit the proteins see only average values of gene expression patterns, averaged over the five gene states. Therefore, in this limit we sum Eqs. (16), (17), (18) to predict the time evolution of N_A, and Eqs. (20) and (21) for the time evolution of N_B to obtain

$$\frac{d\Phi_A}{dt} = \frac{g_{bare}^a + g_{on}^a x_{AA}^a \Phi_A^2 + g_{off}^a x_{BB}^a \Phi_B^2}{1 + x_{AA}^a \Phi_A^2 + x_{BB}^a \Phi_B^2} - \delta^A \Phi_A \tag{24}$$

$$\frac{d\Phi_B}{dt} = \frac{g_{bare}^b + g_{on}^b x_A^b \Phi_A}{1 + x_A^b \Phi_A} - \delta^B \Phi_B, \tag{25}$$

where we have defined $\langle N_S \rangle / N_0 \equiv \Phi_S$, $S = A, B$, and $x_n^m = \frac{h_n^m}{f_n^m}$ with m referring to the respective gene and n indicating the monomer or dimer binding of the transcription factors according to the chosen Hill coefficient in the biochemical reactions.

3.1.1. Comparison with the deterministic description of a bistable frustrated unit

Let us first briefly compare the equations (24,25) with the deterministic equations formerly used to describe the bistable frustrated unit in [8, 15]

$$\frac{d\Phi_A}{dt} = \frac{\alpha}{1 + (\Phi_B/K)} \cdot (\frac{b + \Phi_A^2}{1 + \Phi_A^2}) - \Phi_A \tag{26}$$

$$\frac{d\Phi_B}{dt} = \gamma(\Phi_A - \Phi_B), \tag{27}$$

where γ is the ratio of the half-life of Φ_A to that of Φ_B, that is δ^B/δ^A. In these units, the parameter K sets the strength of repression of Φ_A by Φ_B. The parameter b determines the basal expression level of A, $b < 1$. The parameter α is the maximal rate of production of A for full activation ($\Phi_A^2 \gg b$) and no repression ($\Phi_B \approx 0$). If we divide Eqs. (24 25) by δ^A and define $\tau = t\delta^A$, $\gamma_{on}^a = \frac{g_{on}^a}{\delta^A}$, similarly for the other g-parameters, we have

$$\frac{d\Phi_A}{dt} = \frac{\gamma_{bare}^a + \gamma_{on}^a x_{AA}^a \Phi_A^2 + \gamma_{off}^a x_{BB}^a \Phi_B^2}{1 + x_{AA}^a \Phi_A^2 + x_{BB}^a \Phi_B^2} - \Phi_A \tag{28}$$

$$\frac{d\Phi_B}{dt} = \frac{\delta^B}{\delta^A} \left(\frac{\gamma_{bare}^b + \gamma_{on}^b x_A^b \Phi_A}{1 + x_A^b \Phi_A} - \Phi_B \right). \tag{29}$$

In the previous model we used α as a bifurcation parameter which multiplies Φ_A^2 in the gain term of Eq. (26), a similar role in Eq. (28) is played by g_{on}^a which we therefore use here as the bifurcation parameter. Differently from our former parametrization, apart from the common prefactor δ^B/δ^A, that sets the slow time scale of Φ_B, the gain term in the second equation (29) implicitly depends on $1/\delta^B$, compared to the loss term. Therefore to align the scale of production with the slow decay, we have to adjust the production by choosing $g_{on,bare}^b$ each by two orders of magnitude smaller than the

corresponding production rates of $g^a_{on,bare}$ which explains our choice in table 1. So the slow dynamics of protein B is realized by both slow decay and slow production rate on the genetic level.

Moreover, it should be noticed that we have changed the power of the Hill coefficient in the binding term of the repressor concentration Φ_B from one in Eq. (26) to two in Eq. (28) corresponding to the choice of $h^a_{BB}(i)\frac{N^2_B}{N^2_0}$ in the master equation (4). This appears as a minor difference in the equations. The effect, however, is a considerable broadening of the intermediate limit cycle regime in case of a Hill coefficient of 2. Since we are interested in the fate of the regular oscillations in case of slow and ultra-slow genes, it is important not to need a finetuning for seeing oscillations for fast genes.

Furthermore, we would like to compare our equations (28,29) with the deterministic equations as they were derived for "design I" in [9]. In common with those equations of [9] is the limit of fast genes and the realization of the repression operating on the transcriptional level. The main differences between both sets of equations are firstly the power one of Φ_A in Eq. (29), which can be traced back to the monomer (rather than dimer) binding of the transcription factor A to the promoter region of gene b in Eq. 4, and secondly, the relative weights between gain and loss terms. In particular the bifurcation parameter affects in our case only the first equation directly and the second equation indirectly via Φ_A, while it affects both equations directly in [9]. The combination of these apparently minor differences leads to different bifurcation patterns: a saddle-node bifurcation in [9] and Hopf bifurcations in our case. When increasing our bifurcation parameter g^a_{on}, we see two fixed-point regimes for low and high values of g^a_{on}, separated by an intermediate limit-cycle regime due to two Hopf bifurcations. Therefore, the deterministic equations (28,29) reproduce the phase structure of the bistable frustrated unit. Naively one may expect that the actual bifurcation scenarios in the deterministic limit are irrelevant for the final stochastic systems. It is, however, known from the work of [16] in the context of neural networks and also emphasized in [9] that the very bifurcation scenario may have a strong impact on the final biological function of the motif, as the very onset of oscillations and the embedding in phase space have an impact on amplitude, frequency, noise resistance and other stability properties. Therefore we present a detailed bifurcation analysis of Eqs. (28,29) in the Appendix. As it is seen there, the analysis requires a further zoom into the two transition regions, that is, a high resolution and finetuning of the bifurcation parameter. It would be interesting to search for manifestations or remnants of these scenarios in a fully stochastic description, of which we studied so far the gross features only: "noisy" fixed points and "noisy" limit cycles.

3.1.2. Gillespie simulations for fast genes We present results of Gillespie simulations of the reactions, listed in Eqs. (1) and (2) for parameter values as in Table 1 and the first row of Table 2. Figure 3, left column, shows phase portraits of N_B versus N_A for three values of the bifurcation parameter, $g^a_{on} = 100$ (first row), $g^a_{on} = 300$ (second row) and $g^a_{on} = 900$ (third row), which are typical for the lower ($g^a_{on} = 100$) and higher ($g^a_{on} = 900$)
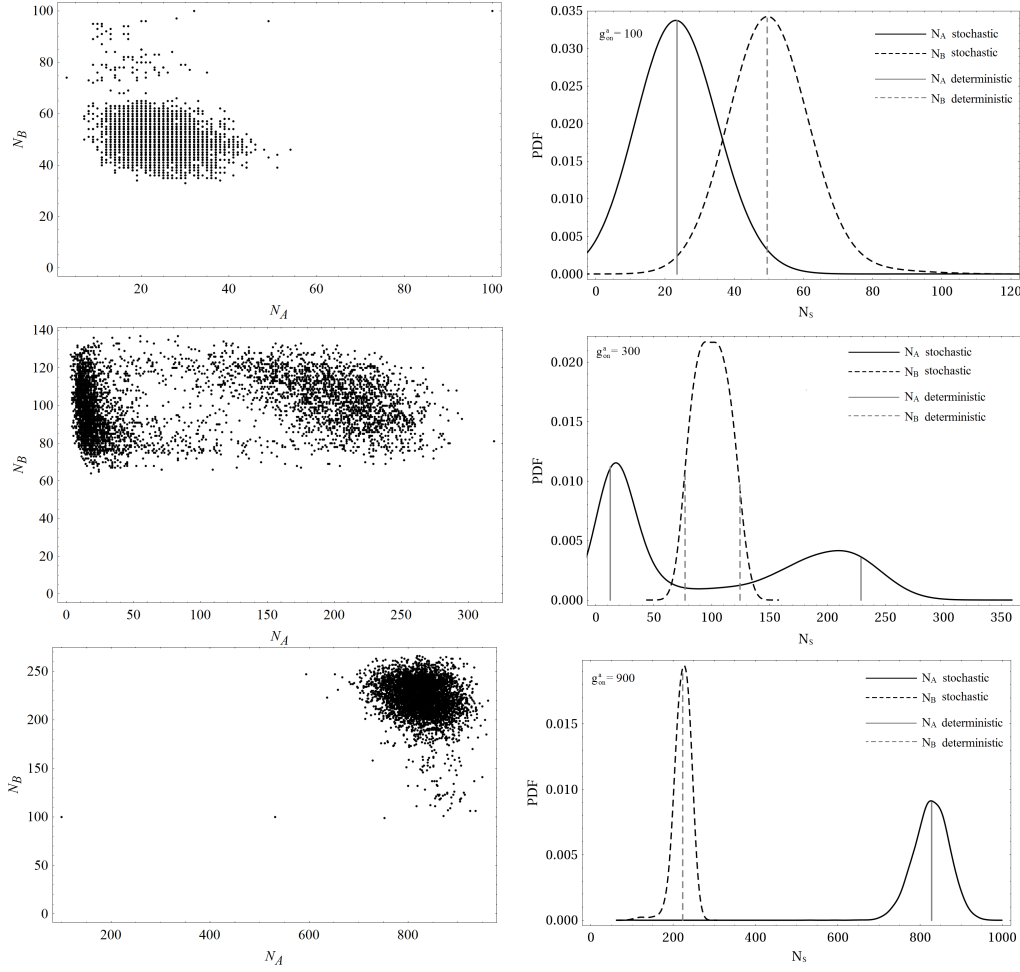
Figure 3. Gillespie simulations for fast genes, that is $h^a_{AA} = h^a_{BB} = 0.01$, $h^b_A = 1.0$, $f^a_{AA} = f^a_{BB} = f^b_A = 100$. Phase portraits of the number of proteins N_B versus N_A within Gillespie time $T_G = 5000$ (left column) and corresponding probability density functions (PDF) (right column) of N_A (full line black) and N_B (dashed line black), while the gray full and dashed vertical lines indicate the position of the fixed points in the first and third row. In the first and third row $g^a_{on} = 100$ and $g^a_{on} = 900$, respectively, in the left panels we see the stochastic pendant of the fixed points observed in the deterministic case. For clarity of the figure we do not plot every Gillespie step, but only 5000 of them. The maxima of the PDFs agree well with the fixed points in the deterministic description. In the second row we see the stochastic version of limit cycles for $g^a_{on} = 300$. The vertical lines here mark the maximal and minimal extension of the limit cycles in Φ_A and Φ_B when integrated as solutions of the deterministic equations (24),(25). These plots confirm our former model (26),(27) as a suitable coarse-grained description.

"fixed-point regime", and for the "limit-cycle regime" ($g^a_{on} = 300$). The phase portraits are made within a Gillespie time of $T_G = \sum_i dt_i = 5000$, where dt_i refers to the time interval, randomly chosen out of a Poisson distribution in the i-th Gillespie update. The regimes are termed after their deterministic pendants: In the deterministic limit ($N_0 \rightarrow \infty$) the clouds of N_A, N_B values in the first and third row would shrink to the
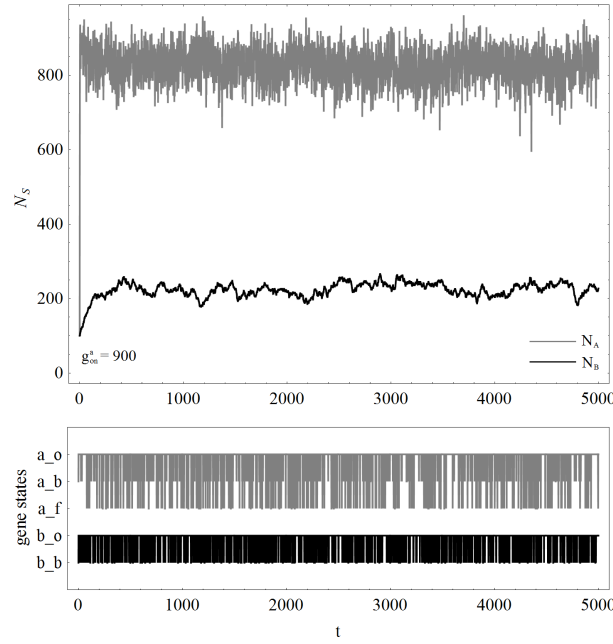
Figure 4. Time series of the number of protein species S, $S = A$ (gray), $S = B$ (black) (upper part of the figure) and the gene states for gene a (gray) and gene b (black) (lower part of the figure), recorded during the Gillespie steps. The bifurcation parameter is chosen as $g_{on}^a = 900$. States of gene a switch between *on*, [a_o (gray)], *bare*, [a_b (white)], *off*, [a_f (gray)], states of gene b switch between *on*, [b_o (black)] and *bare*, [b_b (white)]. The other parameters are chosen as $g_{bare}^a = 25$, $g_{off}^a = 0$, $g_{on}^b = 2.5$, $g_{bare}^b = 0.025$, $h_{AA}^a = h_{BB}^a = 0.01$, $h_A^b = 1$, $f_{AA}^a = f_{BB}^a = f_A^b = 100$, $\delta^A = 1$, $\delta^B = 0.01$.

lower and higher fixed points as predicted from Eqs. (24,25), while the clouds in the second row would contract to a limit cycle. (The higher density of (N_A, N_B)-values for small and large values of N_A is due to the fact that also the stochastic version of a limit cycle spends more time in regions, where N_B drastically changes, since B is the slow variable, while large changes in N_A happen rapidly, since A is the fast variable.) The probability density functions in the right column of the figures reflect the probability for finding concrete combinations of (N_A, N_B)-values in the phase portraits. They are displayed for a quantitative comparison of their maxima with the prediction of the location of the fixed points and the extension of the limit cycle from the deterministic equations. These locations are indicated via the vertical lines. The vertical lines hit the maxima quite well where they correspond to fixed points; the lines also match the typical extension of the cloud in case of the limit cycle.

Figure 4 (upper part) shows the time series of the number of proteins A, N_A (gray) and of proteins B, N_B (black), which are fluctuating about constant values ($N_A \sim 800, N_B \sim 200$). The fluctuations between the different gene states, in particular between the on- and bare-states, is so fast that it appears as a gray and black band (Fig. 4 lower part), so that the protein values, shown in the upper part of the figure, only fluctuate about the fixed point values. In the Gillespie simulations of our former realization of the genetic circuit [8] we identified quasi-cycles deeply in the fixed-point
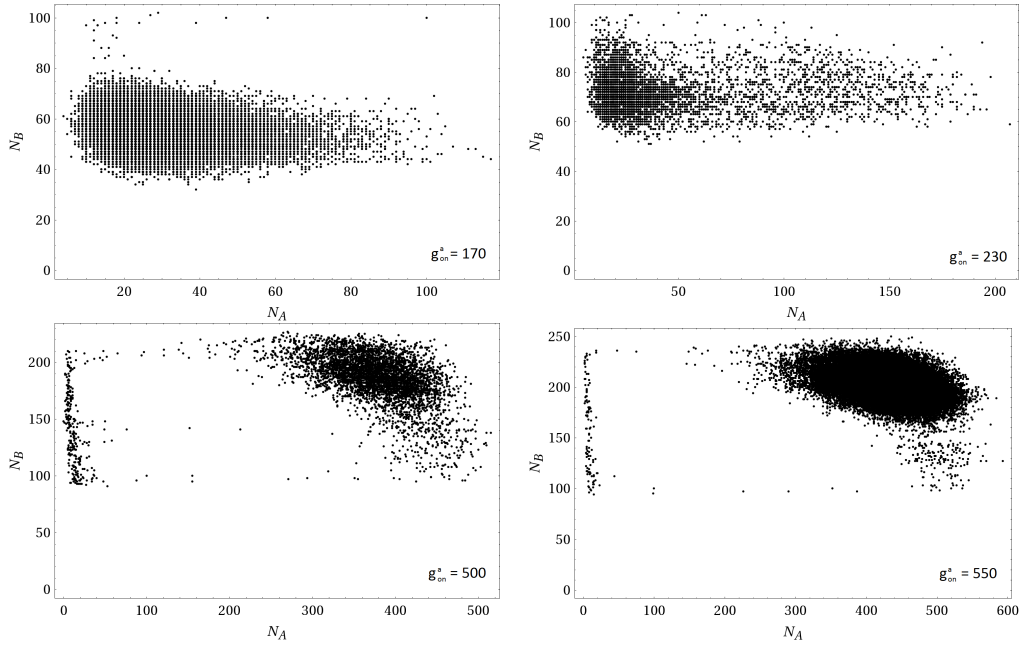
Figure 5. Phase portraits within Gillespie time $T_G = 10000$ for fast genes for four values of the bifurcation parameter. $g_{on}^a = 170$ (upper left), $g_{on}^a = 230$ (upper right), $g_{on}^a = 500$ (lower left), $g_{on}^a = 550$ (lower right). In the lower ($g_{on}^a = 170, 230$) and higher ($g_{on}^a = 500, 550$) fixed point regimes we see clear indications of quasi-cycles in the stochastic descriptions.

regimes. Such cycles, caused by large demographic fluctuations, are also found in the present realization of the genetic circuit if we wait sufficiently long for such fluctuations to happen. Figure 5 shows a few such excursions below (upper left figure for $g_{on}^a = 170$) and above (lower figures for $g_{on}^a = 500$ and 550) the limit-cycle regime, while the data in the upper right figure for $g_{on}^a = 230$ are compatible with more frequent and smaller excursions into phase space due to the vicinity of the transition region to the region of regular limit cycles.

3.2. Slow genes

In the limiting case of slow genes, the binding and unbinding rates are chosen to be of the order of the decay rate of the fast protein, so that they are still fast as compared to the decay rate of the slow protein B. Therefore we sum the moments of N_B from Eqs. (20),(21) also over the states of gene b. Moreover we have three equations (16),(17),(18) that include already a summation over the states of gene b, but in order to further sum over states of gene a, we consider Fig. 6, which shows the switching events of gene a between the three states *on, bare, off* and of gene b between the two states *on* and *bare*. The switching events between the on- and the bare-state, determined by h_{AA}^a, f_{AA}^a, as well as the switching between the *bare-* and the *off*-state, determined by h_{BB}^a, f_{BB}^a, are still frequent as compared to an indirect switching between an *on-* and an *off*-state of gene a, since it is due to different binding events. Therefore the proteins A and B effectively see
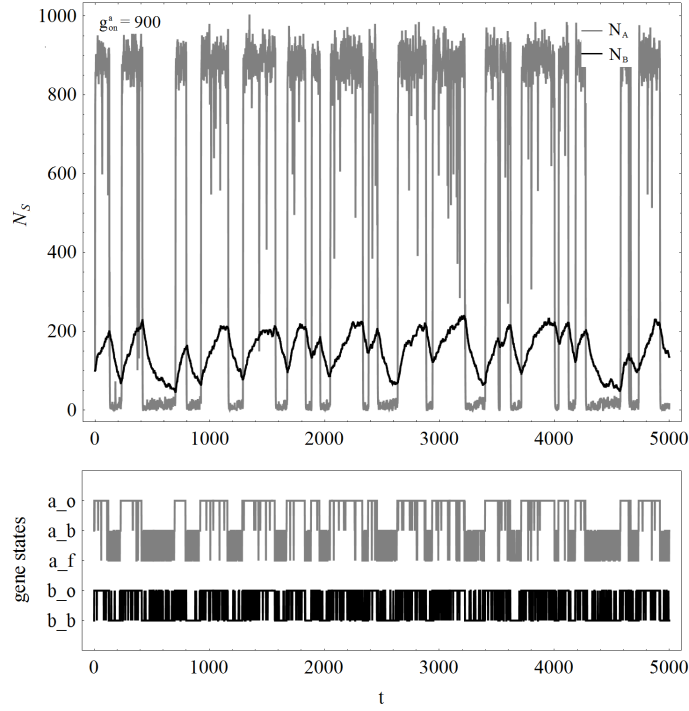
Figure 6. Same as Fig. 4, but for slow genes, that is $h_{AA}^a = h_{BB}^a = 0.0001$, $h_A^b = 0.01$, $f_{AA}^a = f_{BB}^a = f_A^b = 1.0$. Here the switching between *on* and *bare* states is much faster than the indirect switching between *on* and *off* states. The other parameters are as in Fig. 4.

two gene levels, an average over the on- and bare-state, or over the bare- and off-state. These two different levels are seen in the time evolution of N_A: N_A switches between the upper gray band in Fig. 6, that is between $N_A = 800$ and $N_A = 1000$, and the lower gray band close to zero. The adaptation of N_B to the abrupt changes in N_A is more smooth and happens with delay. (As we shall later see, the fast proteins A are able to adapt to these two states, while the slow ones B are not.) Accordingly we average Eq. (16) with Eq. (18) and Eq. (17) with Eq. (18) to obtain two sets of differential equations

$$\frac{d\langle N_A \rangle}{dt} = N_0 \frac{g_{bare}^a + g_{on}^a x_{AA}^a \frac{\langle N_A \rangle^2}{N_0^2}}{1 + x_{AA}^a \frac{\langle N_A \rangle^2}{N_0^2}} - \delta^A \langle N_A \rangle, \tag{30}$$

$$\frac{d\langle N_B \rangle}{dt} = N_0 \frac{g_{bare}^b + g_{on}^b x_A^b \frac{\langle N_A \rangle}{N_0}}{1 + x_A^b \frac{\langle N_A \rangle}{N_0}} - \delta^B \langle N_B \rangle \tag{31}$$

and

$$\frac{d\langle N_A \rangle}{dt} = N_0 \frac{g_{bare}^a + g_{off}^a x_{BB}^a \frac{\langle N_B^2 \rangle}{N_0^2}}{1 + x_{BB}^a \frac{\langle N_B^2 \rangle}{N_0^2}} - \delta^A \langle N_A \rangle \tag{32}$$

$$\frac{d\langle N_B \rangle}{dt} = N_0 \frac{g_{bare}^b + g_{on}^b x_A^b \frac{\langle N_A \rangle}{N_0}}{1 + x_A^b \frac{\langle N_A \rangle}{N_0}} - \delta^B \langle N_B \rangle, \tag{33}$$

which describe the two alternatives of the time evolution of the system, either following the first or the second set of equations. While the time dependence of $\langle N_B \rangle$ is the same in

both equations (31) and (33), the time dependence of $\langle N_A \rangle$ depends on whether it is due to the activating binding of a dimer of A or the repressing binding of B. This apparently minor detail leads to a very different bifurcation pattern, as it is revealed by a linear stability analysis. For Eqs. (32),(33) this analysis shows that the determinant of the Jacobian, evaluated at the single fixed point, is positive for all combinations of positive values of parameters and $g_{bare}^a > g_{off}^a$ and $g_{on}^b > g_{bare}^b$, while the trace of the Jacobian, evaluated at the fixed point, is negative, so the fixed point is always stable. In contrast, for the first set of equations, Eqs. (30),(31), we have for our usual choice of parameters (used in the Gillespie simulations) a single fixed point, but there is a combination of parameters for a very low value of g_{bare}^a, for which we obtain three positive real fixed points; for example choosing $g_{on}^a = 300$, $g_{bare}^a = 5$ and the other parameters as before, we obtain two stable nodes with positive determinant, negative trace and negative eigenvalues at $(N_A^* = 6.09, N_B^* = 16.71)$, and at $(N_A^* = 262.65, N_B^* = 181.75)$, and a saddle with negative determinant, positive trace, and $\lambda_1 < 0$, $\lambda_2 > 0$ at $(N_A^* = 31.26, N_B^* = 61.44)$. Depending on the initial conditions, the system will evolve in one of the stable nodes. Gillespie simulations for this parameter set are not displayed in Fig. 7 below. The basin of attraction of the lower fixed point is so small that the stochastic system practically never evolves to this fixed point due to the demographic fluctuations.

3.2.1. Gillespie simulations for slow genes While for fast (ultra-slow) genes, gene switches are fast (slow) with respect to both proteins, for slow genes protein A sees slow switches between the $on - bare$ and $off - bare$ average values, so that it can follow these different states of gene a, while protein B is intrinsically so slow that it still sees averages over all states of the genes. Therefore inserting the separate average values of $\langle N_A \rangle$ in Eqs. (31),(33) fails as an effective description. The phase portraits of N_A, N_B for slow genes in the left column of Fig. 7 shows the pendant of one fixed point of Eq. (32) (upper row), while the fixed point of Eq. (30) is not visible (two fixed points each in the middle and lower row of the figure), where the left cloud corresponds to the solution of Eq. (32) and the right cloud to Eq. (30). In particular we interpret the clouds of events in the second row also as a stochastic switching between two fixed points rather than a noisy version of limit cycles, since in contrast to the corresponding phase portrait in Fig. 3 there is no empty space between large and small N_A values, the events jump from the left-to the right side rather than performing full cycles, in agreement with the time series of N_A in Fig. 6. The vertical lines in Fig. 7 indicate the two stable fixed points of Eqs. (30),(32) in all three cases. The second fixed point of Eq. (30) is not visible for $g_{on}^a = 100$ (first row), but it is seen for $g_{on}^a = 300$ and $g_{on}^a = 900$, where both fixed point locations fit well the maxima of the PDFs for N_A, though not for N_B. The failure is due to the inherent difficulty that protein B sees averages over all b-states, leading to values of N_B which are not solutions of Eqs. (31),(33) (derived under the condition on protein A to see N_A states as two distinct averages over ($on - bare$) and $off - bare$) states). B is still too inert to follow the two distinct $\langle N_A \rangle$-values.
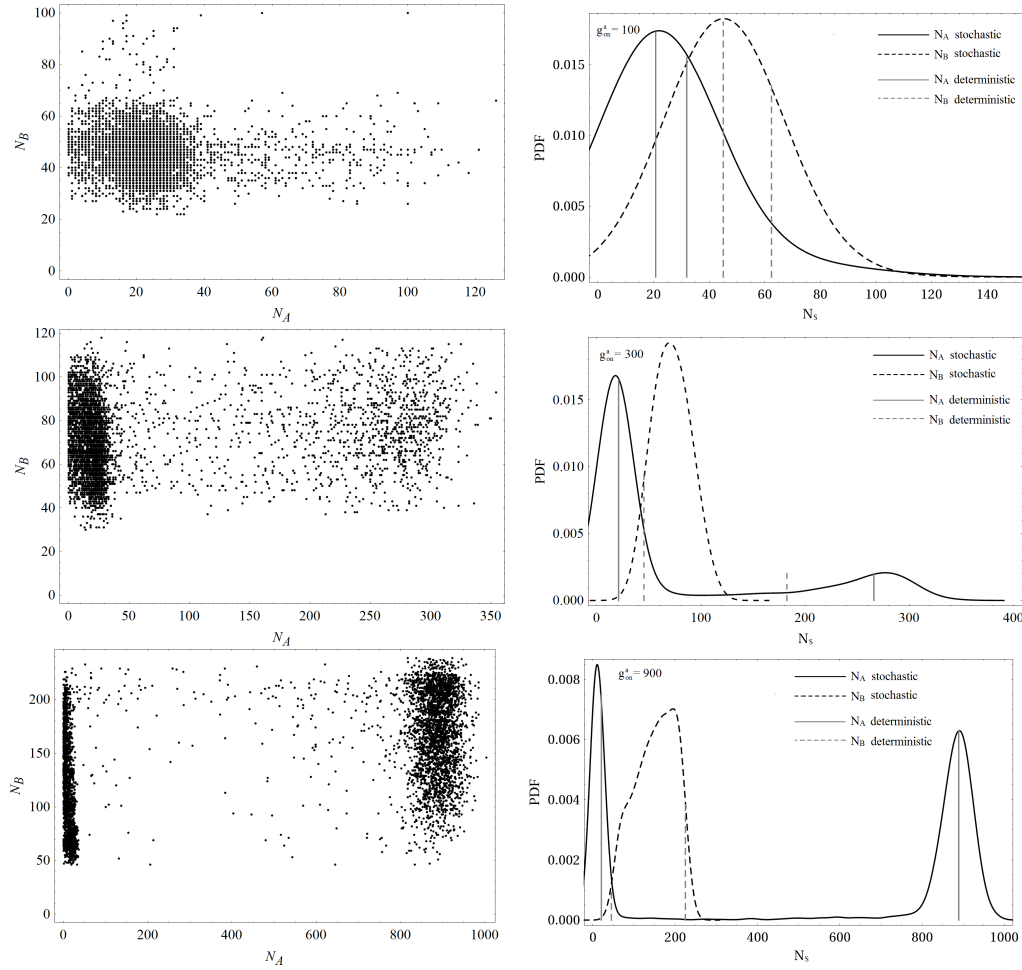
Figure 7. Same as Fig. 3, but for slow genes, that is $h_{AA}^a = h_{BB}^a = 0.0001$, $h_A^b = 0.01$, $f_{AA}^a = f_{BB}^a = f_A^b = 1.0$. The two vertical lines in each species in the PDFs indicate the prediction of two stable fixed points. This is in agreement with the stochastic results only for $g_{on}^a = 900$ (third row). For $g_{on}^a = 100$ (first row) we see the stochastic pendant of a single fixed point only in disagreement with the prediction from the deterministic model. The interpretation of the phase portrait in the second row remains ambiguous: the second local maximum in the PDF seems to be a precursor of the second fixed point, and the interpretation of the phase portrait in terms of a limit cycle is less likely due to the absence of a white area which should not be visited by the Gillespie trajectory in case of a limit cycle. For further explanations we refer to the main text.

3.3. Ultra-slow genes

Independently of the possibility to realize this limit in natural or synthetic genetic circuits, it is of interest from the dynamical point of view what the effective coarse-grained description in the deterministic limit amounts to in case of ultra-slow genes. This limit refers to a situation, in which the binding/unbinding rates of transcription factors are of the order of the slow protein B. So the time which genes a and b spend in one of their possible states is long as compared to $1/\delta^A$, the lifetime of protein A. It then does no longer make sense to further sum Eqs. (16)–(18) over any states of gene a
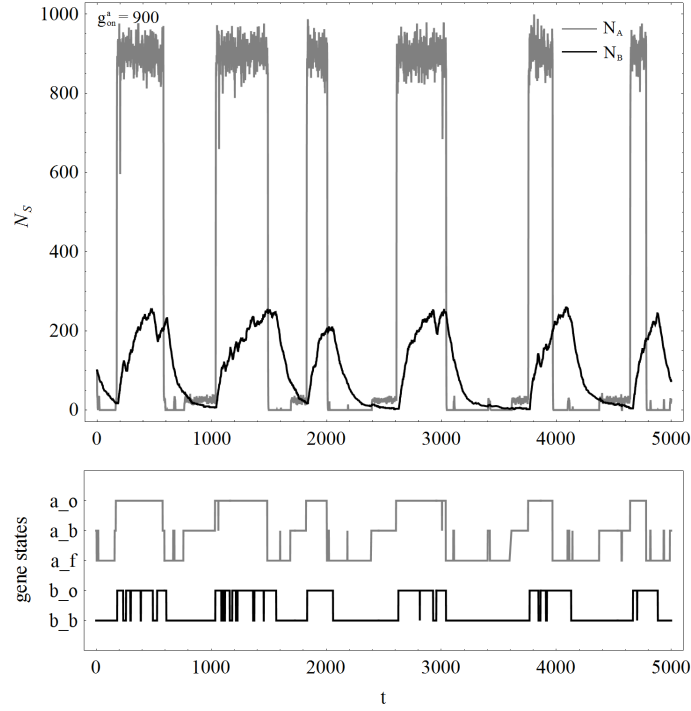
Figure 8. Same as Fig. 4, but for ultra-slow genes, that is $h^a_{AA} = h^a_{BB} = 0.000005$, $h^b_A = 0.0001$, $f^a_{AA} = f^a_{BB} = f^b_A = 0.01$. Protein A sees gene a in three different states, while there are periods during which protein B still sees averages over *on* and *bare*-states, here visible in the first two blocks, where gene b switches frequently between *on* and *bare*.

and Eqs. (20),(21) over any states of gene b, but we are left with these five equations, which can be summarized as

$$\frac{d(\langle N_S \rangle s_i)}{dt} = s_i \frac{d\langle N_S \rangle}{dt} + \langle N_S \rangle \frac{ds_i}{dt} \tag{34}$$

with $S = A, B$, and $s_i = a_{on}, a_{bare}, a_{off}$ for $S = A$ and b_{on}, b_{bare} for $S = B$, respectively. If we insert for ds_i/dt Eq. (22) and solve (34) for $d\langle N_S \rangle/dt$, we arrive at the following set of uncoupled differential equations for $\langle N_S \rangle/N_0 =: \Phi_S$

$$\frac{d\Phi_A}{dt} = g^a_i - \delta^A \Phi_A , \quad i = on,\ bare,\ off$$

$$\frac{d\Phi_B}{dt} = g^b_i - \delta^A \Phi_A , \quad i = on,\ bare \tag{35}$$

with solutions that for $t \to \infty$ exponentially decay to the fixed points g^s_i/δ^S with $S = A, B, s = a, b, i = on, bare, off$ for $s = a$, and $i = on, bare$ for $s = b$, leading to six fixed points.

3.3.1. Gillespie simulations for ultra-slow genes   In the stochastic realization of this limit of ultra-slow genes we expect the system to switch between three possible states with respect to N_A and two with respect to N_B, so between six fixed points in the deterministic limit. The former oscillations in the limit cycle regime are clearly gone. For
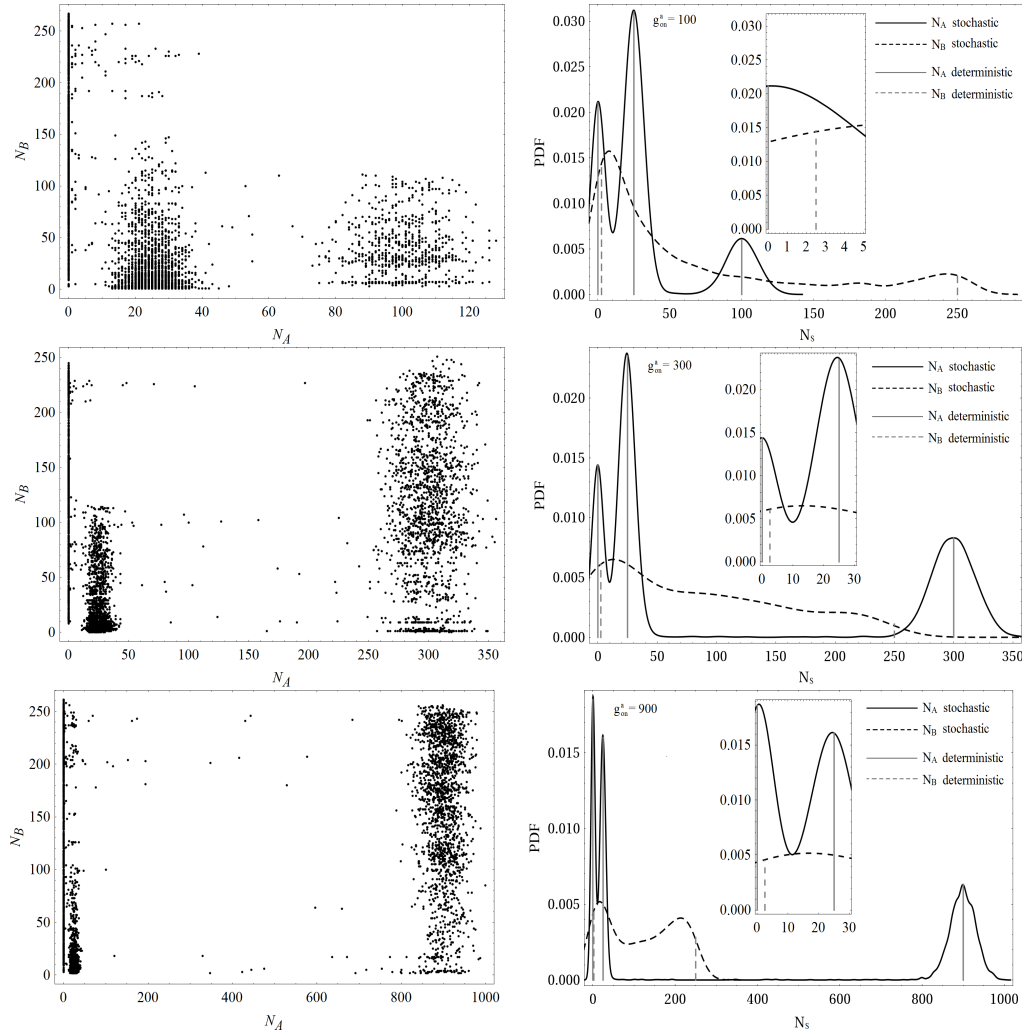
Figure 9. Same as Fig. 3, but for ultra-slow genes, that is that is $h^a_{AA} = h^a_{BB} = 0.000005$, $h^b_A = 0.0001$, $f^a_{AA} = f^a_{BB} = f^b_A = 0.01$. In all three rows, left column, we see remnants of three fixed points in the deterministic limit with respect to values of N_A, while the values of N_B are broadly spread, since B is too slow to follow the different states of gene b. The vertical lines from the deterministic prediction of the fixed point values (right column) match well the maxima of the PDFs for N_A and only roughly for N_B. The insets zoom into the (N_A, N_B) values of the two fixed points for lower N_A-values.

N_A we see both in the phase portraits and in the probability density functions remnants of three distinct fixed-point values of N_A, while the remnants of two possible fixed-point values of N_B are only vaguely visible as two maxima in the probability distribution. Obviously the ultra-slow genes are still not slow enough to allow protein B to adjust to the different states of gene b.

4. Summary and conclusions

The motif of the self-activating species which activates its own repressor is found in many realizations in biological systems. From the physics' point of view one would like to identify universal features of the dynamical behavior. Certainly regimes of excitable and oscillatory behavior are common features found over a wide range of parameters. Also in our current realization we have recovered three regimes of excitable, oscillatory and excitable behavior as function of one bifurcation parameter, which we have chosen as g_{on}^a, the production rate of protein A if gene a is in the on-state. However, as we have pointed out, these three distinct regimes are only obtained in our realization if the binding and unbinding rates of genes are fast as compared to the other inherent time scales, here the decay rates of the fast (A) and the slow (B) proteins. For this case we derived deterministic equations equivalent to the former ones of [15]. As soon as the binding/unbinding rates are no longer small, but of the same order as the decay time of either proteins, the averaging procedure for deriving a deterministic limit has to be changed; the proteins see no longer average values of the gene states, but tend to follow the distinct states, unless their own production is too slow to reach the appropriate state "in time". These features were demonstrated by our Gillespie simulations of the six master equations. For the ultra-slow genes they were also visible in the derived coarse-grained description in terms of five deterministic equations. For the intermediate case of binding rates of the order of the decay of the fast protein, the intrinsic difficulty to derive deterministic equations lies in the fact that one and the same binding and unbinding events are differently seen from the proteins: protein A sees already distinct events and follows the according gene state, protein B still sees an average over the gene states. In both cases, slow and ultra-slow genes, the intermediate regime of stable regular oscillations is absent.

One should keep in mind that we have to deal with systems of nonlinear dynamics. Therefore one should be aware that apparently minor changes such as the value of the Hill coefficient may have pronounced effects. Even if the qualitative picture after such a minor change remains the same, the quantitative features like the extension of the limit-cycle regime can drastically change, as we have seen. Such a change can finally decide on the relevance of the model for the real biological system. For real systems stable oscillations would probably not be observed if they were restricted to a tiny interval of the bifurcation parameter.

Although we have chosen our parameters independently of their possible realization in concrete genetic circuits, the conclusion from our analysis is generic and applies also to real biological systems: It is not only the gross features of the topology of the motifs and the couplings that determine the dynamical performance. If different time scales are inherent, the gross bifurcation patterns may depend on their ratios.
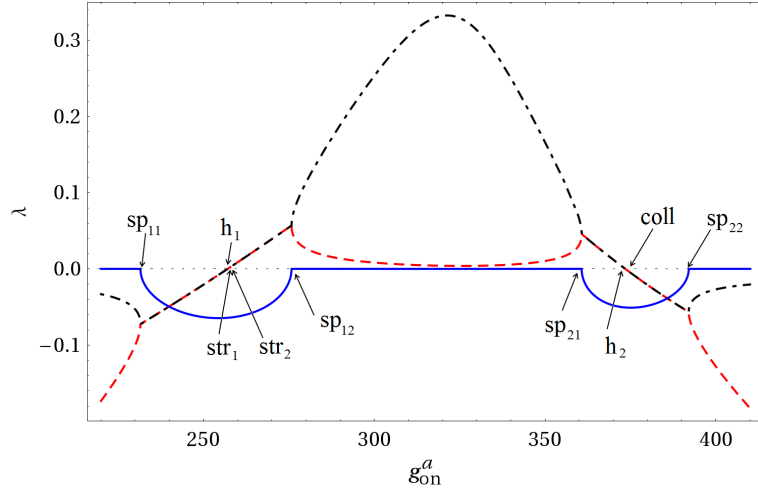
Figure 10. Real (black and red (gray)) and imaginary (blue) part of the eigenvalues of the Jacobian as function of the bifurcation parameter g_{on}^a. To keep the figure clear we present only the negative imaginary part. For the meaning of the various indicated special values of g_{on}^a we refer to the main text.

Acknowledgment

Appendix: Zoom into the bifurcations of the deterministic equations for fast genes

The appendix is devoted to a detailed bifurcation analysis of Eqs. (24), (25). In Fig. 10 we plot the real-and imaginary part of the two eigenvalues of the Jacobian $\partial \dot{\Phi}_S / \partial \Phi_{S'}, S, S' \in \{A, B\}$, evaluated at the respective fixed points of these equations. We display only the negative imaginary part to keep the figure clear. In the following we describe the stationary states as a function of increasing parameter g_{on}^a. The labels assigned to the specific values of g_{on}^a shall remind to the change in the dynamical performance occurring at these values (so that spirals are created at "sp", Hopf bifurcations at "h", two limit cycles collide at "coll" and "str" marks the boundaries of the tiny "strange" interval in which two limit cycles coexist).

- $g_{on}^a < sp_{11} = 231.5492$: one stable node.
  In this interval we observe one fixed point in the form of a stable node.
- $sp_{11} = 231.5492 \leq g_{on}^a < h_1 = 257.0829$: one stable spiral.
  At $sp_{11} = 231.5492$ the two real eigenvalues become complex conjugate of each other, and the stable node turns into a stable spiral, shown in Fig. 11.
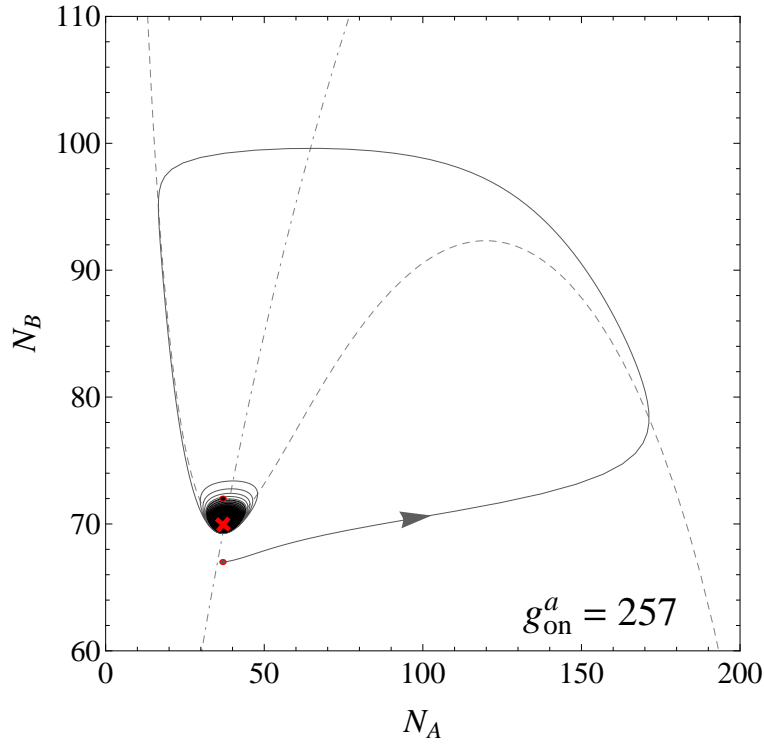
Figure 11. Stable spiral. Starting in the vicinity of the fixed point (red (gray) cross) either from above or from below (both starting points indicated by red dots), the trajectory spirals into the fixed point, either directly, when starting from above, or after a long excursion in phase space, starting from below. The other parameters are chosen as $g_{bare}^a = 25$, $g_{off}^a = 0$, $g_{on}^b = 2.5$, $g_{bare}^b = 0.025$, $h_{AA}^a = h_{BB}^a = 0.01$, $h_A^b = 1$, $f_{AA}^a = f_{BB}^a = f_A^b = 100$, $\delta^A = 1$, $\delta^B = 0.01$. Gray dashed and dashed-dotted lines show the nullclines of N_A and N_B.

- $h_1 = 257.0829 \leq g_{on}^a < str_1 = 257.12233$: one unstable spiral and one small stable limit cycle.

  At the Hopf bifurcation point $h_1 = 257.0829$ the real part of the eigenvalues crosses the abscissa and becomes positive, so that the stable spiral becomes unstable. From that value on, as long as $g_{on}^a < str_1$, there is a small stable limit cycle with a radius growing with increasing g_{on}^a, surrounding the unstable spiral. Trajectories spiral out of the fixed point towards the limit cycle and spiral extremely slowly when approaching the limit cycle, cf. Fig. 12. The first Lyapunov coefficient $l_1(h_1) = -3.751804 \cdot 10^{-6}$, where the negative sign of l_1 shows that the Hopf bifurcation is supercritical.

  Without a further zoom into the bifurcation parameter, one would conclude from the observation of a large stable limit cycle for $g_{on}^a > str_2 = 257.12253$ (see below) that the small stable limit cycle for $g_{on}^a \geq h_1$ has continuously grown to this large size. What happens instead is described next.

- $str_1 = 257.12233 \leq g_{on}^a < str_2 = 257.12253$: two coexisting stable limit cycles and one unstable spiral.
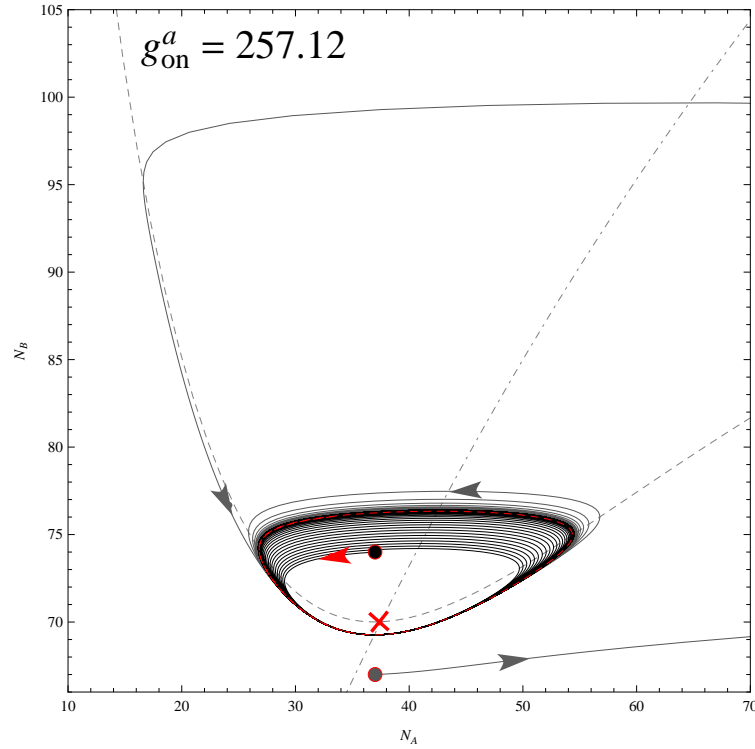
Figure 12. Unstable spiral and small stable limit cycle. Starting now in the vicinity of the fixed point (red (gray) cross) from above or below, the trajectory ends up in a small stable limit cycle indicated as red dashed line. Parameters other than g_{on}^a are chosen as in Fig. 11.

Figure 13 shows a snapshot in bifurcation parameter space, just after the second limit cycle is born. The two black stripes in the inset consists of two limit cycles, a smaller and a larger one, and the white space in between is filled by the trajectories, independently of the initial conditions, as long as this space is sufficiently small.

The second limit cycle pops up just after the small stable limit cycle from Fig. 12 has grown until $g_{on}^a = str_1$. The phase space between the smaller and the larger limit cycle acts like an attractor as it is seen in Fig. 14: trajectories, starting from initial conditions in this intermediate space (two initial conditions are indicated by the gray and black dot) follow limit cycles in the area between the small and the large limit cycle, indicated as gray dashed lines in the lower right part of the figure, see Fig. 14. The zoom presented in the lower and upper right part shows two time snapshots up to time $T = 2000$ and $T = 10000$, respectively, where T denotes the total sum over elementary time intervals.

The area between the limit cycles rapidly grows (Fig. 15 left), until only one large stable limit cycle remains (Fig. 15 right).

- $str_2 = 257.12253 \leq g_{on}^a < sp_{12} = 257.7795$: one large stable limit cycle and one unstable spiral.

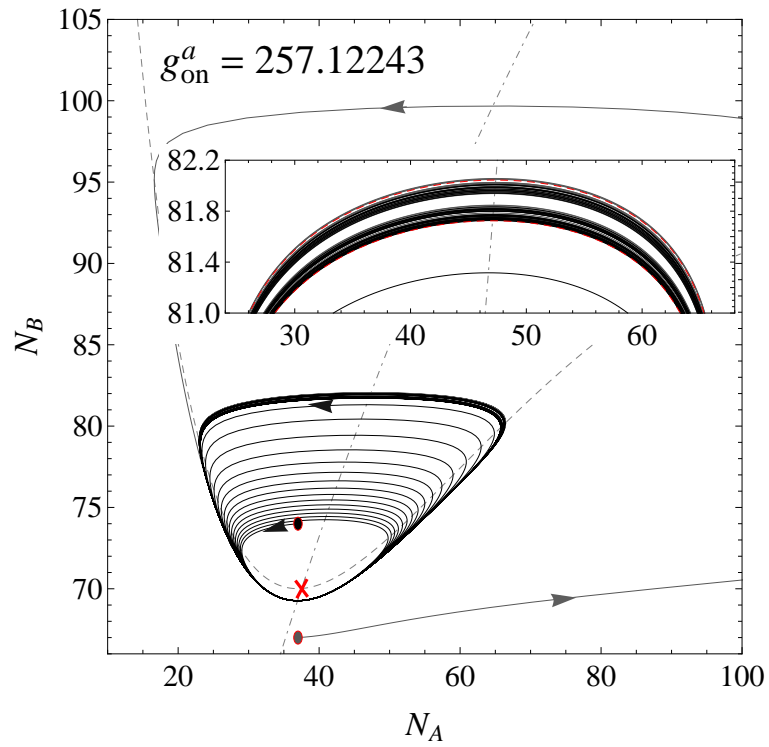 From $g_{on}^a = str_2$ on only one large stable limit cycle is left along with the unstable

Figure 13. An instant in bifurcation parameter space, shortly after a second stable limit cycle is born. The coexistence of two stable limit cycles is visible as two black bands in the inset. Parameters other than g_{on}^a are chosen as in Fig. 11.
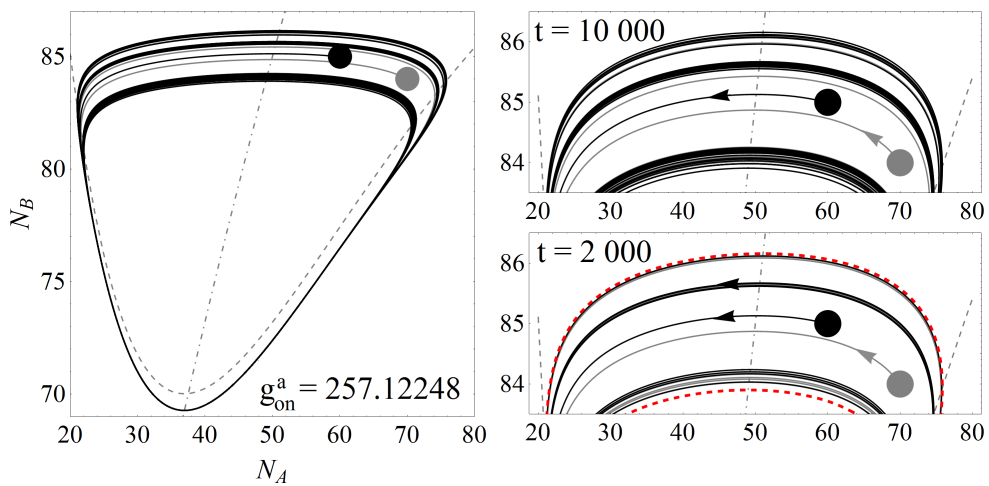

Figure 14. Two coexisting limit cycles for $g_{on}^a = 257.12248$ and other parameters chosen as in Fig. 11. Trajectories starting between the lower and upper limit cycle remain in this area bounded by the red (gray) dashed lines in the lower right part of the figure.
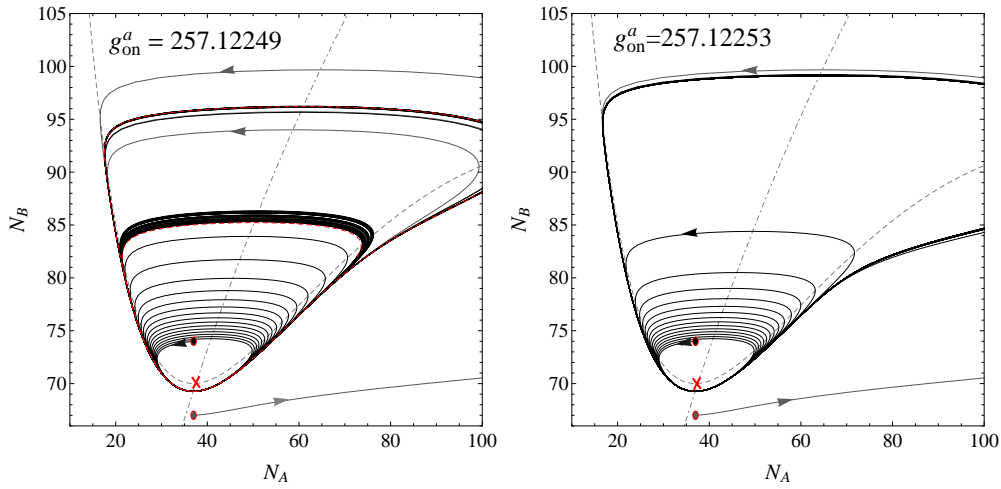
Figure 15. Left: Independently of the starting point, the trajectory will first hit the closer one of both limit cycles, then it will not stay exactly on this cycle, but continue in its vicinity. Right: From this value of g_{on}^a on, only one large limit cycle survives.

fixed point which stays a spiral until $g_{on}^a = sp_{12}$.

- $sp_{12} = 257.7795 \leq g_{on}^a < sp_{21} = 360.71422$: one large stable limit cycle and one unstable node.
  At $g_{on}^a = sp_{12}$, the imaginary part of the complex conjugate eigenvalues become zero, leaving two real different positive eigenvalues, so that the fixed point inside the large stable limit cycle has turned into an unstable node. This situation holds over a large interval in g_{on}^a, until the unstable node turns back into an unstable spiral, not displayed.

- $sp_{21} = 360.71422 \leq g_{on}^a < h_2 = 373.4836$: one large stable limit cycle and one unstable spiral.

  The return to an unstable spiral happens at $g_{on}^a = sp_{21}$, where the eigenvalues become again complex conjugate. The spiral remains unstable until the second Hopf bifurcation happens.

- $h_2 = 373.4836 \leq g_{on}^a < coll = 373.90582$: one stable spiral, one small unstable limit cycle and one large stable limit cycle.
  The second Hopf bifurcation happens at $g_{on}^a = h_2$, where the real part of the complex conjugate eigenvalues crosses the abscissa, the formerly unstable spiral changes into a stable one. Along with that a small unstable limit cycle is born, while the large stable limit cycle is still "alive". So we have a stable spiral, surrounded by a small unstable limit cycle, surrounded by the large stable limit cycle (not visible in Fig. 16), see Fig. 16.
  This coexistence holds only for a tiny interval, as long as $g_{on}^a < coll$. The reason for the small size of this interval is that the fixed point and the small unstable limit cycle are located close to the trajectory of the large stable limit cycle, so that a tiny increase of g_{on}^a leads to a growth of the radius of the small unstable cycle that is
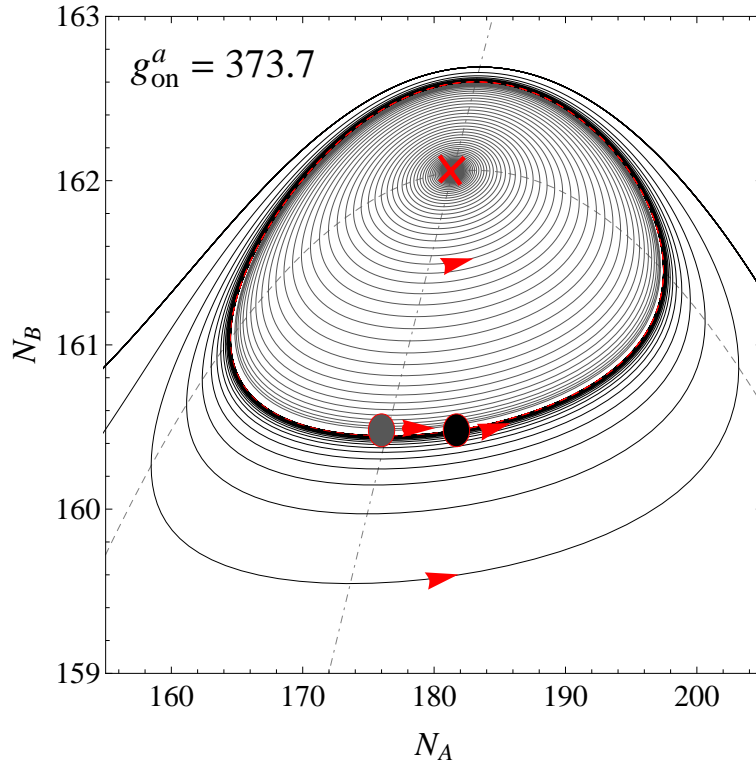
Figure 16. Two starting points inside and outside the unstable limit cycle lead to a spiral into the fixed point and to the large limit cycle (not displayed), respectively. Parameters other than g_{on}^a are chosen as in Fig. 11.

sufficient for the small cycle to collide with the large cycle and this way to induce its cancelation. The collision of the two limit cycles amounts to a saddle-node bifurcation of two periodic orbits with the result that the periodic orbits disappear. This suggests the vicinity of a Bautin bifurcation [17]; the Bautin bifurcation is characterized by two bifurcation conditions: (i) $Re\lambda_{1,2} = 0$ and (ii) the critical first Lyapunov coefficient $l_1(\alpha_{\mathrm{crit}}) = 0$. We have in our interval $h_2 \leq g_{on}^a < coll$ that $Re\lambda_{1,2} = 0$ and $l_1(h_2) = 9.5 \cdot 10^{-6}$ at the Hopf bifurcation point.

- $coll = 373.90582 \leq g_{on}^a < sp_{22} = 392.0632$: one stable spiral, no limit cycle anymore. After the stable and unstable limit cycles disappear at $g_{on}^a = coll$, there remains a stable spiral as long as $g_{on}^a < sp_{22}$.

- $sp_{22} = 392.0632 \leq g_{on}^a$: one stable node.
  At $g_{on}^a = sp_{22}$ the two complex conjugate eigenvalues become real again with two different negative values, so that finally the stable spiral returns to a stable node and remains like that for increasing g_{on}^a.

While the linear stability analysis reveals the change from a stable spiral to an unstable spiral, and an unstable node back to an unstable spiral, to a stable spiral and a stable node, the numerical integration of the equations (24),(25) shows subtleties in how the large stable limit cycle is born and destroyed. The creation happens discontinuously and in coexistence with a small stable limit cycle, the annihilation very likely via a Bautin bifurcation.

References

[1] J. Martiel, A. Goldbeter, Biophys. J. **52**, 807 (1987).

[2] B. Novak, J. J. Tyson, J. Cell Sci. **106**, 1153 (1993).

[3] J. R. Pomerening, S. Y. Kim, J. E. Jr. Ferrell, Cell **122**, 565 (2005).

[4] J. J. Tyson, Proc. Natl. Acad. Sci. U.S.A. **88**, 7328 (1991).

[5] L. Qiao, R. B. Nachbar, I. G. Kevrekidis, S. Y. Shvartsman, PLoS Computational Biology **3**, e184 (2007).

[6] J. M. G. Vilar, H. Y. Kueh, N. Barkai, S. Leibler, Proc. Natl. Acad. Sci. U.S.A.**99**, 5988 (2002).

[7] N. T. Ingolia, A. W. Murray, Current Biology **14**, R771 (2004).

[8] A. Garai, B. Waclaw, H. Nagel and H. Meyer-Ortmanns, J. Stat. Mech. **P01009** (2012).

[9] R. Guantes, J. F. Poyatos, PLoS Comp. Biol. **2**, 0188 (2006).

[10] R. FitzHugh, Biophys. J. **1**, 445 (1961).

[11] J. S. Nagumo, S. Arimoto, S. Yoshizawa, Proc. IRE **50**, 2061 (1962).

[12] A. M. Walczak, M. Sasai, P. G. Wolynes, Biophys. J. **88**, 828 (2005).

[13] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[14] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, (North Holland, Amsterdam, 2005).

[15] P. Kaluza, H. Meyer-Ortmanns, Chaos **20**, 043111 (2010).

[16] E. M. Izhikevich, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 1171 (2000).

[17] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, (Springer, Heidelberg, 1998).